MODELING, SIMULATION AND APPLICATION OF BACTERIAL TRANSDUCTION IN GENETIC ALGORITHMS

Carlos Perales-Gravan, Javier de Vicente Buendia, Juan Castellanos,R. Lahoz-Beltra

***Abstract**: At present, all methods in Evolutionary Computation are bioinspired by the fundamental principles of neo-Darwinism, as well as by a vertical gene transfer. Virus transduction is one of the key mechanisms of horizontal gene propagation in microorganisms (e.g. bacteria). In the present paper, we model and simulate a transduction operator, exploring the possible role and usefulness of transduction in a genetic algorithm. The genetic algorithm including transduction has been named PETRI (abbreviation of Promoting Evolution Through Reiterated Infection). Our results showed how PETRI approaches higher fitness values as transduction probability comes close to 100%. The conclusion is that transduction improves the performance of a genetic algorithm, assuming a population divided among several sub-populations or 'bacterial colonies'.*

***Keywords:** Bacterial genetic algorithm, horizontal gene transfer, conjugation operator, transduction operator.*

***ACM Classification Keywords**:* I.6 SIMULATION AND MODELING

Introduction

At present, all methods in Evolutionary Computation (genetic algorithms, evolutive algorithms, genetic programming, etc.) are *bioinspired* by the fundamental principles of neo-Darwinism [Lahoz-Beltra, 2008), and by a vertical gene transfer; that is to say, by a mechanism in which an organism receives genetic material from the ancestor from which it evolved. Indeed, most thinking in Evolutionary Computation focuses upon vertical gene transfer as well as upon crossover and/or mutation operations.

Microorganisms have been evolving on Earth for billions of years. At present, there are several reasons why microbial evolution experiments have been attracting increasing attention [Elena and Lenski, 2003). Microbiologists have long known how bacteria are capable of adapting and evolving in all kinds of environments. Bacteria are microscopic organisms whose single cells reproduce by means of a process of binary fission or of asexual reproduction, bearing a resemblance to John von Neumann's universal constructor [von Neumann, 1966]. Thus, a bacterial population (or colony) evolves according to an evolutive algorithm similar to Dawkin's biomorphs [Dawkins, 1986], the cumulative selection of mutations powering their evolution. Bacteria, however, exhibit significant phenomena of genetic transfer and crossover between cells. This kind of mechanism belongs to a particular kind of genetic transfer known as horizontal gene transfer. Horizontal, lateral or cross-population gene transfer is any process in which an organism, i.e. a donor bacterium, transfers a genetic segment to another one, a recipient bacterium, which is not its offspring. In the realm of biology, whereas the scope of vertical gene transfer is the population, in horizontal gene transfer the scope is the biosphere. This particular mode of parasexuality between 'relative bacteria' includes three genetic mechanisms: conjugation, transduction and transformation. In a previous paper [Perales-Gravan and Lahoz-Beltra, 2008], we introduced an evolutionary algorithm through the substitution of crossover operations in a genetic algorithm by means of conjugation. Furthermore, microorganisms are very interesting individuals because they also exhibit 'social interactions'.

Recently we found [Lahoz-Beltra et al., 2009] how the inclusion of the 'social life of microorganisms' into the genetic algorithm cycle, significantly improves the algorithm's performance.

In Nature, microorganisms such as bacteria and viruses share a long and common evolutionary relationship. This relationship is mainly promoted by bacteriophages (or phages) [Davis et al., 1990], a kind of virus that multiplies inside bacteria by making use of the bacterial biosynthetic machinery. Some bacteriophages are capable of moving bacterial DNA (the 'bacterial chromosome') from one bacterium to another. This process is known as transduction. When bacteriophages infect a bacterial cell, their normal mode of reproduction makes use of the bacterium's replication machinery, making numerous copies of its own viral genetic material (i.e. DNA or RNA). The nucleic acid copies (or chromosome segments) are then promptly packaged into newly synthesized copies of bacteriophage virions. Considering the life cycle of a particular bacteriophage, we can define two sorts of transduction (Figure 1). Generalized transduction occurs when 'any part' of the bacterial chromosome (rather than viral DNA) hitchhikes into the virus (i.e. T4 phages in *Escherichia coli* bacterium). However, when only 'specific genes' or certain special 'segments' of the bacterial chromosome can be transduced, such a mistake is known as specialized transduction (i.e. λ phages in *Escherichia coli* bacterium). Here we study the possibility of developing genetic algorithms, including transduction operations as a horizontal gene transfer mechanism. The efficiency and performance of transduction was evaluated using a benchmark function and the 0/1 knapsack problem. The utility was illustrated by designing an AM radio receiver and optimizing the main features of the electronic components of the AM radio circuit, as well as those of the radio enclosure. Our results show how transduction improves the performance of a genetic algorithm, assuming a population divided among several sub-populations or 'bacterial colonies'. Consequently, in transduction, transference of chromosome segments between bacterial populations or colonies is very different from migration (the occasional exchange of individuals). Migration and transduction could bear a resemblance, but only when transduction involves the complete chromosome transference between bacterial populations. Furthermore, this kind of transference is a highly unlikely event in bacteria, transduction of chromosome segments taking place in these microorganisms.

In this paper, we model and simulate the two kinds of transduction operations examining the possible role and usefulness of this genetic mechanism in genetic algorithms. In a previous paper [Perales-Gravan and Lahoz-Beltra, 2008], we introduced a bacterial conjugation operator showing its utility by designing an AM radio receiver. Conjugation is one of the key genetic mechanisms of horizontal gene transfer between bacteria. In the present paper, we refer to a genetic algorithm including transduction as PETRI (*Promoting Evolution Through Reiterated Infection*). We investigated the transfer of genes and chromosomes among sub-populations with a simulated 'bacteriophage'. In the model we consider a structured population divided among several sub-populations or 'bacterial colonies', bearing a resemblance with coarse-grain distributed genetic algorithms. Each sub-population is represented as a Petri dish (a glass or plastic cylindrical dish used to culture microorganisms). It should be noted, however, that even when we divide a population into sub-populations, the proposed algorithm is sequential. Thus, the algorithm is not a distributed one, since we used a mono-processor computer and the algorithm was not parallelized. Moreover, the migration mechanism is synchronous, as gene and chromosome transferences were both between sub-populations and during the same generation. Therefore, our approach could be related with those models of Cellular Genetic Algorithms (cGA) adopted also for mono-processor machines [Alba and Dorronsoro, 2008], with no relation to parallelism at all. In our model, we assumed that bacteria are capable of displaying crossover through conjugation, instead of performing one-point or two-point recombination. Moreover, we assume that no vertical gene transfer mechanism is present in bacterial populations.

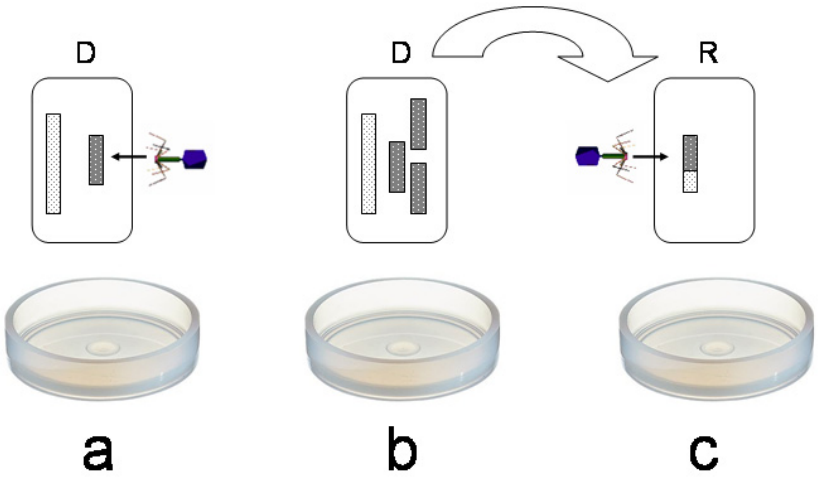

Figure 1. Transduction mechanism (bacterium DNA, white rectangle; bacteriophage DNA, grey rectangle). (a) Infection of a donor bacterium D with a bacteriophage. (b) Bacterial and bacteriophage DNA segments mix inside donor bacterium D. A bacterial DNA segment is packed inside the bacteriophage 'head' (c), and is transferred to a recipient bacterium R. Finally, inside the recipient bacterium, R homologous recombination or crossover occurs between the emigrant bacterial DNA segment and the target bacterial chromosome.

With the aim of studying the performance of the transduction operator, we used different optimization problems. Experiments conducted in the presence of transduction were compared with control experiments, performed in the absence of transduction. Similarly, we compared the transduction performance under the three types of crossover: conjugation, one-point or two-point recombination. We are interested in the study of genetic algorithms based on horizontal gene transfer mechanisms, mainly conjugation and transduction operations. It is important to note that even when conjugation and transduction are both horizontal gene transfer mechanisms, there are some relevant differences between both. In the first place, whereas conjugation involves two bacteria from the same population, the bacteria involved in transduction can belong to different populations. As a consequence, conjugation is a genetic mechanism of horizontal gene transfer *within* a population, whereas transduction is a genetic mechanism of horizontal gene transfer *between* populations. Secondly, in conjugation, the length of the transferred genetic segment is variable, whereas in transduction, the transferred segment length is always constant.

Model description

In this section, we introduce the transduction model, as well as the PETRI implementation, that is, the algorithm that results once transduction is included in a genetic algorithm.

Transduction model

Let *b* be a chromosome (i.e. bacterium; 1,.., *j* ,..., *N*) and *p* a sub-population (i.e. Petri dish; 1 ,..., *i* ,..., *P*); then a transduction operation (Figure 2) is defined as follows: transduction is the transfer of genetic material from a Petri

dish and bacterium donors (p^D, b^D) to a Petri dish and bacterium recipients (p^R, b^R). When the transference involves a chromosome segment, the result is a recombinant chromosome in the recipient Petri dish p^R. However, the transference of a complete chromosome results in the substitution of one chromosome of the recipient Petri dish p^R with the transferred one. It is important to note that 'bacterium' and 'Petri dish' terms are used throughout the paper as 'chromosome' and 'sub-population' synonyms, respectively. Transduction requires the selection of the Petri dish and bacterium donors (p^D, b^D), as well as the Petri dish and bacterium recipients (p^R, b^R) In the following site **http://bioinformatica.net/tests/petri.html** we describe how transduction was conducted.

PETRI: A genetic algorithm with simulated transduction

The current PETRI (*Promoting Evolution Through Reiterated Infection*) algorithm (Figure 2) uses a population size of *N*, performing r_e replicates, with *P* being the total number of Petri dishes or sub-populations. Thus, we performed a number of $r_e.P$ trials of each simulation experiment. The algorithm cycles through epochs, searching for an optimum solution until a maximum of *G* generations is reached. Once (p^D, b^D) and (p^R, b^R) are selected, only one 'bacteriophage' is assumed to participate during each transduction event. The PETRI algorithm is summarized in the following pseudocode description:

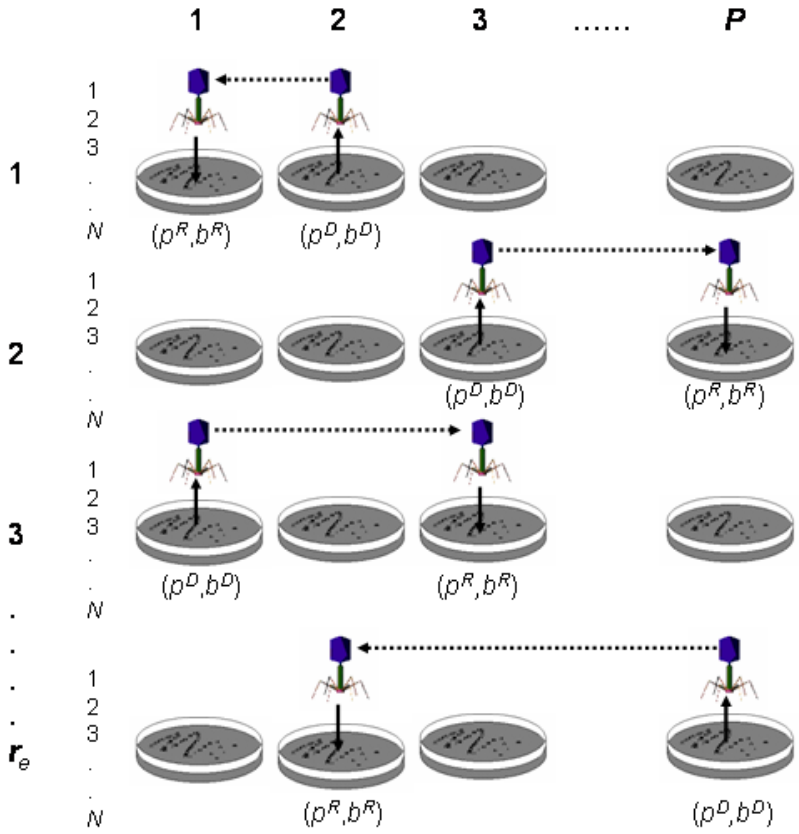

Figure 2. Transduction experiment. The figure shows transduction from donor Petri dish (*p*D) and bacterium (*b*D) to recipient Petri dish (*p*R) and bacterium (b^R). In the figure, *P* is the total number of Petri dishes (or sub-populations), *N* is the number of bacteria (or population size) per Petri dish and r_e the number of experimental replicates.

```
/* PETRI: Genetic Algorithm with Transduction */
        1.      t:=0;
        2.      Initialization: Generate P Petri dishes (or sub-populations) with N random bacteria (or chromosomes).
        3.      WHILE not stop condition DO
            /* Genetic Algorithm */
                        (3.1)  FOR each P Petri dish DO
                                    Evaluation of chromosomes
                                    Selection
                                    Conjugation or Crossover  (one-point, two-point)
                                    Mutation
                               (3.2)  END FOR
                    /* End of Genetic Algorithm */
        4.      Transduction: (pD, bD) → (pR, bR)
        5.      t:=t+1;
        6.      END WHILE;
/* End of PETRI */
```

Starting out with a random population of chromosomes (or bacteria), we simulated selection, crossover, mutation, and transduction, obtaining new generations of equal population size. Once the initial population of chromosomes was randomly obtained, the order in which the genetic operators were applied was in consonance with the SDS protocol [Lahoz-Beltra, 2001] [Perales-Gravan and Lahoz-Beltra, 2004]. This protocol is inspired by DNA shuffling, an experimental method used in biotechnology for improving *in vitro* protein activity and functionality. In the present simulation experiments, the SDS protocol involves a cycle of crossover and mutation, as well as transduction, through n_g generations. This phase is followed by repeated cycles of crossover and transduction, one per generation, in absence of mutation. Considering the performance of previous experiments, the simulation experiments were conducted setting n_g equal to 25. Selection was simulated as follows. In each generation, we evaluated the fitness of each chromosome using a fitness function that depends on the optimization problem chosen. Once the chromosomes were evaluated, we selected the crossover (or mating) pool of the next generation using the roulette wheel parent selection algorithm [Goldberg, 1989] [Lahoz-Beltra, 2004]. Clearly, other selection schemes are possible, such as tournament selection, truncation selection, as well as linear and exponential ranking selection. However, the roulette wheel parent selection scheme bears a better resemblance to Darwinian natural selection [Lahoz-Beltra, 2001]. Once a new generation of offspring chromosomes is obtained, then pairs of chromosomes are randomly selected *within* a sub-population or Petri dish. Once a pair of chromosomes (or bacteria) {#*i*, #*j*} is selected, whether or not to perform crossover on the current pair of chromosomes {#*i*, #*j*} is decided on the basis of a Bernoulli trial regarding conjugation as having a given probability p_c (or alternatively, instead of conjugation, crossover is conducted via one-point or two-point recombination). We simulated mutation of a gene by randomly changing the value gene, choosing the mutated value from a uniform distribution with a similar range to those defined to obtain the initial population of chromosomes. Once again, whether or not to change a gene value on a chromosome is decided on the basis of a Bernoulli trial, mutation being a success with a given probability p_m (mutation probability). Transduction was simulated based on the model described above. The operator requires the selection of the Petri dish and bacterium donors (p^D, b^D), as well as the Petri dish and bacterium recipients (p^R, b^R). Once (p^D, b^D) and (p^R, b^R) are both selected, whether or not to perform transduction on the current pair is decided on the basis of a Bernoulli trial regarding transduction as having a given probability p_t (transduction probability).

Simulation experiments

We studied the performance of the simulated transduction by considering three optimization problems. The first problem uses a benchmark function, the second one is the 0/1 knapsack problem, and finally we illustrated the usefulness of transduction in the problem described in [Perales-Gravan and Lahoz-Beltra, 2008]:

Experiment 1. - An initial optimization problem was an instance of the Michalewicz function. We used 10 variables or genes, so that:

$$f(x) = \sum_{i=1}^{10} sen(x_i) \cdot \left(sen\left(\frac{i \cdot x_i^2}{\pi} \right) \right)^{2m}; \qquad 0 \le x_i \le \pi \tag{1}$$

The simulation experiments were conducted with N = 500 chromosomes (or bacteria), r_e = 15 and P = 4 Petri dishes. Thus, we performed 60 trials (15 replicates x 4 Petri dishes) in each experiment. The transduction experiments were conducted transferring only complete chromosomes, cycling the algorithm through epochs searching for an optimum, until a maximum of 700 generations (G) is reached. In each trial, we calculated the population average fitness in the last generation. The crossover and mutation probabilities were set to p_c=0.75 and p_m=0.05, respectively. When crossover was simulated through conjugation, then the conjugation parameter was set as $\alpha = 0.5$.

Since preliminary results (*experiment 3*) indicated the best transduction policy, we simulated transduction by selecting donors (p^D, b^D) based on *max-max* criterion, and recipients (p^R, b^R) using *roul-r* criterion. The simulation experiments were conducted by setting the transduction probabilities p_t to 0% (control experiment, without transduction), 25%, 50%, 75% and 100%. We performed simulation experiments with PETRI, using conjugation, and one-and two-point recombination.

Experiment 2. - A second optimization problem was the well-known 0/1 knapsack problem. Let us assume we have j kinds of items and each item has a value v_j and weight w_j, the maximum weight that we can carry in the knapsack being equal to W. The 0/1 knapsack problem restricts the number of each kind of item x_j to 0 or 1. The aim is to maximize $\sum_j v_j x_j$ subjected to $\sum_j w_j x_j \le W$. Fitness was calculated by the usual expression:

$$f(x) = \begin{cases} \sum_j w_j x_j \le W, \sum_j v_j x_j \\ \\ \sum_j w_j x_j > W, W - \sum_j w_j x_j \end{cases} \tag{2}$$

using the benchmark knapsack instance 'knap100' published on a web site by the [Swiss Federal Institute of Technology Zurich, 2008]. The instance includes the values and weights of 100 items, maximum weight being W=2732. The simulation experiments were performed with N = 200 chromosomes (or bacteria), r_e = 15 and P = 4 Petri dishes, conducting 60 trials (15 replicates x 4 Petri dishes) in each experiment. The transduction experiments transferred only complete chromosomes, cycling the algorithm through epochs until a maximum of 1000 generations (G) was reached. In each trial, we obtained maximum fitness in the last generation. The crossover and mutation probabilities were set to p_c=0.75 and p_m=0.05, respectively. When crossover was simulated through conjugation, the conjugation parameter was set as $\alpha = 0.5$. Once again, transduction was

simulated by selecting donors (p^D, b^D) based on the *max-max* criterion and recipients (p^R, b^R) with the *roul-r* criterion. The transduction probability p_t was set to 0% (control experiment, without transduction), 25%, 50%, 75% and 100%. We performed simulation experiments with PETRI, using conjugation and one- and two-point recombination.

Experiment 3. - An example of the usefulness of transduction in a real-life problem consists of finding the main features of the electronic components of an AM radio receiver, along with those of the radio enclosure. The current PETRI algorithm uses a population size of 500 (N) and 9 (P) Petri dishes, performing fifty replicates (r_e). Therefore, we performed 450 trials (50 replicates x 9 Petri dishes) in each experiment. The algorithm cycles through epochs, searching for an optimum AM radio receiver until a maximum of 200 generations (G) is reached. In each trial, we calculated the population average fitness in the last generation. The crossover and mutation probabilities were set to p_c=0.75 and p_m=0.05, respectively. As crossover was simulated via conjugation, the parameter α was set to 0.5. The initial population of chromosomes was obtained at random, choosing the gene values from a uniform distribution according to the ranges described in [Perales-Gravan and Lahoz-Beltra, 2008]. In each generation, we evaluated the fitness of each chromosome, that is, the degree of achievement of the AM radio receiver circuit, along with the main features of the radio enclosure. Considering the intricacy of the fitness evaluation, for a detailed explanation, see [Perales-Gravan and Lahoz-Beltra, 2008].

The simulation experiments performance was evaluated according the statistical methods described in [Lahoz-Beltra and Perales-Gravan, 2010] as well as using the MAF values defined in the site http://bioinformatica.net/tests/petri.html.

Note that figures cited below (in section Results) are included in the aforementioned web site.

Results

A remarkable result showing the role of transduction was obtained with the Michalewicz function. Figure 3 shows how PETRI approaches the maximum function value (9.66 in our experiments) as transduction probability p_t comes close to 100%. A Kruskal-Wallis test shows that, with a p-value equal to zero, the differences among medians were statistically significant at the 95.0% confidence level. Thus, the genetic algorithm performance is significantly improved, regardless of the crossover operator (conjugation, one- or two-point recombination). In particular, in the absence of transduction, the medians of conjugation, one-point recombination and two-points recombination were 8.87, 8.96 and 8.97, respectively. However, when optimization experiments included a transduction operator (transferring a complete chromosome), the medians of conjugation, one-point recombination and two-points recombination were 9.29, 9.33 and 9.36, respectively. The Bartlett, Cochran, and Levene tests were accomplished, testing the homogeneity of variances. The obtained p-values were zero, showing that differences among variances are statistically significant at the 95.0% confidence level. The conclusion is that variance (or population variability) depends upon transduction. Indeed, the population variability reaches a minimum value when a complete chromosome is transferred and the transduction probability p_t equal to 100%. However, it is important to note that in Nature, only tiny fragments or chromosomal segments of bacterial DNA are transduced, as opposed to complete chromosomes. In agreement with [Davis and Weller, 1998], the genetic material carried by bacteriophages is, conveniently, around 2% the length of the bacterial chromosome. An interesting observation is that even when transduction is a major driving force behind diversity in natural populations, in our simulation experiments, transduction reduces the variability of the population. The explanation might be that, in natural populations, transduction occurs at random and at very low frequencies,

from 10^{-2} to 10^{-10} [Ogunseitan, 2008], whereas in the experiments performed, the chromosome or its segments were selected with a medium or high probability and based on *max-max* criterion.

Figure 4 shows a Multiple-Box-and-Whisker Plot of the maximum fitness values obtained in the 0/1 knapsack problem. It is interesting to note how PETRI approaches highest fitness values as transduction probability p_t comes close to 100%. A Kruskal-Wallis test shows that, with a *p*-value equal to zero, the differences among medians were statistically significant at the 95.0% confidence level. Consequently, algorithm performance is significantly improved regardless of the crossover operator employed (conjugation, one-point recombination or two-point recombination). In the absence of transduction, the median value of the obtained knapsacks under conjugation, one-point recombination and two-point recombination were 2601.5, 2639.5 and 2612.0 respectively. However, in presence of transduction, with a p_t equal to 100% (transferring a complete chromosome), the median values of the optimized knapsacks were 3214.0, 3253.0 and 3126.0, respectively. The Bartlett, Cochran, and Levene tests were accomplished to examine the homogeneity of variances. The obtained *p*-values were zero, showing that differences in variances between the set of experiments without transduction (labeled as 1, 6 and 11 in Figure 4) and the experiments including transduction are statistically significant at the 95.0% confidence level. Once again, we conclude that variance depends on transduction, population variability decreasing when a complete chromosome is transferred.

Based upon the above results, we reached the following general conclusion. Transduction helps the population to reach a better optimum solution, and has an effect on population variability. However, the optimum achieved with the transduction of chromosome segments is always below in relation to the optimum reached when transduction transfers a complete chromosome. On the other hand, when transduction involves chromosome segments, population variability (or variance) is greater than the variability that results from a complete chromosome transduction. That is, in consonance with Figure 6, transduction of complete chromosomes, a mechanism that bears a resemblance to migration (e.g. animals and plants), will push a population to the highest optimum but to the lowest variability. In contrast, transduction of chromosome segments, a mechanism that is closer to real transduction in microorganisms, for instance, will move the population toward a higher variability, but a lower optimum value (Figure 6). Both situations could represent different strategies of organisms during evolution, preventing premature convergence [Grefenstette, 1981]. Transduction of chromosome segments results in a subsequent homologous recombination or crossover, promoting the sudden jump of the recipient population towards better solutions in the evolutive or fitness landscape. This might be explained by the fact that the arrival and recombination of new genetic information (including good Holland's schemata) breaks the population equilibrium (Gould's punctuated equilibrium; see [Gould and Eldredge, 1977]) in the recipient Petri dish, triggering its evolutionary change [Cohoon et al., 1987]. Our findings support similar observations made in Nature. For instance, by evaluating the coevolution of the bacterium *E. coli* and the bacteriophage T7, [Forde et al., 2004] found that the local adaptation was lower in closed communities than in open ones, suggesting that gene flow was acting as source of beneficial mutations in the open communities.

Figure 5 shows the Multiple Box-and-Whisker Plot of the MAF_T values obtained for each one of the thirty-three types of transduction simulation experiments performed with the AM radio receiver (*experiment 3*). The Kruskal-Wallis test shows that, with a *p*-value below 2.200×10^{-16}, there are statistically significant differences among the medians at the 95.0% confidence level. Comparing the control experiment (the simulation experiment without transduction, labeled as 1 in Figure 5) with the two best transduction experiments (a chromosome segment is transferred, labeled as 19 in Figure 5, and the transference of a complete chromosome, labeled as 33 in Figure 5), we concluded as follows. Regardless of the kind of transduction, the best results were obtained (Figure 6) when selection of the donor Petri dish and the bacterium (p^D, b^D) are both based on the *max-max* criterion, the recipients (p^R, b^R) being selected with the *roul-r* criterion. Note that the best transduction protocol is the one in

which the simulated 'bacteriophage' selects both the donor Petri dish and the bacterium with maximum fitness, so that $\max_{1,..,P}\left\{\overline{f}_1,\overline{f}_2,...,\overline{f}_i\right\}$ and $\max_{1,..,N}\left\{f_1,f_2,...,f_j\right\}$, whereas selection of the recipient Petri dish and the bacterium are both based on stochastic methods, such as the roulette wheel approach or the uniform distribution procedure. In the Kruskal-Wallis test between the control experiment and the transduction experiments with chromosome segments (Figure 6), the *p*-value obtained was equal to 1.603×10^{-08}. Since the *p*-value is below 0.05, there is a statistically significant difference between the medians at the 95.0% confidence level. This is similar to the Kruskal-Wallis test between the control experiment and the transduction experiments with complete chromosomes (Figure 6). In this case, the *p*-value obtained was below 2.200×10^{-16}, the differences between the medians at the 95.0% confidence level being statistically significant.

Figure 7a shows a representative performance graph (a Box-and-Whisker Plot per generation) of the control experiments (without transduction and r_e=50), as well as the graph obtained for the best transduction experiments in which only chromosome segments are transferred (Figure 7b) and the transduction experiment transfers a complete chromosome (Figure 7c). Furthermore, Figure 8 shows in the aforementioned experiments the mean of the average fitness per generation (MAF). In the control experiments (Figure 8a), the MAF_T is equivalent to the MAF $\overline{D}\,\overline{R}$ value per generation. Note how in transduction experiments the highest slope (Figure 8b-8c) corresponds to the donor Petri dishes (MAF_D). In such figures, note the overlapping between the performance curves, that is the curve of the recipient Petri dishes (MAF_R) and the curve of the MAF value per generation calculated for all the Petri dishes (MAF_T). The oscillating behavior of MAF_R is explained by the fact that the Petri dish and the bacterium recipients are both selected by means of stochastic procedures. Obviously, the worst performance is observed in those Petri dishes or colonies that do not participate in the transduction experiments (MAF $\overline{D}\,\overline{R}$).

Discussion

Thirty years ago, [Anderson, 1970] suggested that 'virus transduction' could be considered as one of the key mechanisms of horizontal gene propagation. This fact suggests the importance of horizontal gene transfer as an evolutionary mechanism, as it involves the transport of DNA segments from individuals belonging to one phylum to individuals of another phylum. Furthermore, the evolutionary dynamics of populations might depend on the transfer of DNA from one population to other. In fact, [Syvanen, 1985] suggests that cross-species gene transfer could help to explain many experimental observations. For instance, the rapid bursts in evolution (Gould's punctuated equilibrium hypothesis, see [Gould and Eldredge, 1977]), or the widespread occurrence of parallelism (or convergent evolution of similar traits in the fossil record). According to [Margulis, 1981], the acquisition and accumulation of random mutations is not sufficient to explain how inherited variations occur. Moreover, whereas Darwinism emphasizes the role of selection as the force behind evolution (indeed, selection is the main evolutive mechanism underlying genetic algorithms), [Margulis, 1981] and other evolutionary biologists emphasize the role of horizontal gene transfer and cooperation.

The simulation results are consistent with the general picture of transduction in bacteria. Transduction and conjugation operators sufficiently capture the role of such microbial genetic mechanisms in horizontal gene transfer. PETRI efficiency is shown by optimizing a benchmark function, solving the 0/1 knapsack problem and evolving an optimum design of an AM radio receiver. Even when PETRI is not a distributed genetic algorithm, most DGA and cGA applications use a simple process of chromosome migration as the only mechanism of horizontal gene transfer. However, we show how, based upon microbial genetics, it is possible to develop new

algorithms, genetic operators and simulation protocols, which might be of use in Evolutionary Computation. For instance, in PETRI algorithm we combine two different genetic mechanisms exhibited by real bacteria. Of a side conjugation, that is, a local mechanism of horizontal information transfers *within* a population; of other transduction, that is, a global mechanism of horizontal information transfers *between* populations.

In a previous paper, [Kubota et al., 1996] introduced VEGA, addressing the possibility of simulating transduction, a virus-evolutionary genetic algorithm. Their authors introduced a virus infection operator and a virus fitness value, modeling two populations, the host population and the virus population. The main difference between this model and our model is the fact that, in VEGA, there are two populations: a host population, representing the candidate solutions, and a virus population (initially generated from host population), representing a substring set of solutions. Consequently, in VEGA, the underlying model is the coevolution of two populations, a main host population and a secondary viral population. In contrast, in PETRI several hosts or bacterial populations (or Petri dishes) share solutions via transduction. Furthermore, VEGA viruses propagate their own substrings or chromosome segments among host individuals, whereas in PETRI viruses propagate only the host (or bacterial) chromosome or segments between the host individuals. Another interesting difference is that VEGA viruses represent a real population, each virus presenting a fitness value. However, our model viruses are 'dummy' agents, responsible for the transduction mechanism. Finally, in VEGA, the host population does not comprise bacteria, as crossover is simulated in the usual manner in genetic algorithms that is one-point or two-point recombination (vertical gene transfer). In contrast, PETRI simulates crossover based on a conjugation operator (horizontal gene transfer) or alternatively, as VEGA, thus using one-point or two-point recombination (vertical gene transfer).

Conclusion

We modeled and simulated a transduction operator, exploring the usefulness of transduction in a genetic algorithm. The genetic algorithm including transduction has been named PETRI, showing how PETRI approaches higher fitness values as transduction probability comes close to 100%. The conclusion is that transduction improves the performance of a genetic algorithm, assuming a population divided among several sub-populations or 'bacterial colonies'.

Bibliography

[Alba and Dorronsoro, 2008] E. Alba, B. Dorronsoro, B. 2008. Cellular Genetic Algorithms. Springer.

[Anderson, 1970] N.G. Anderson. 1970. Evolutionary significance of virus infection. Nature 227: 1346-1347.

[Cohoon et al., 1987] J.P. Cohoon, W.N. Martin, D.S. Richards. 1987. Punctuated equilibrium: A parallel genetic algorithm. In: Proceedings of the Fifth International Conference on Genetic Algorithms (Ed. J.J. Grefenstette). Lawrence Erlbaum Associates, Hillsdale, NJ: 148-154.

[Davis and Weller, 1998] R.H. Davis, S.G. Weller. 1998. The Gist of Genetics. Guide to Learning and Review. Jones & Bartlett Publishers, Sudbury, MA.

[Davis et al., 1990] B. D. Davis, R. Dulbecco, H.N. Eisen, H.S. Ginsberg. 1990. Microbiology. Lippincot Company, Philadelphia.

[Dawkins, 1986] R. Dawkins. 1986. The Blind Watch maker. W. W. Norton & Company, Inc, New York, NY.

[Elena and Lenski, 2003] S.F. Elena, R.E. Lenski. 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. Nature Genetics 4: 457-469.

[Forde et al., 2004] S.E. Forde, J.N. Thompson, B.J.M. Bohannan. 2004. Adaption varies through space and time in a coevolving host-parasitoid interaction. Nature 431: 841-844.

[Goldberg, 1989] D.E. Goldberg. 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, Reading, MA.

[Gould and Eldredge, 1977] S.J. Gould, N. Eldredge. 1977. Punctuated equilibria: the tempo and mode of evolution reconsidered. Paleobiology 3: 115-151.

[Grefenstette, 1981] J.J. Grefenstette. 1981. Parallel adaptive algorithms for function optimization. Tech. Rep. No. CS-81-19, Vanderbilt University, Computer Science Department, Nashville, TN.

[Kubota et al., 1996] Kubota, N., Shimojima, K., Fukuda, T. (1996). The role of virus infection in virus-evolutionary genetic algorithm. In: Proceedings of the IEEE International Conference on Evolutionary Computation, Nagoya, Japan: 182-187

[Lahoz-Beltra and Perales-Gravan, 2010] R. Lahoz-Beltra, C. Perales-Gravan. 2010. A survey of nonparametric tests for the statistical analysis of evolutionary computational experiments. International Journal Information Theories and Applications 17(1): 49-61.

[Lahoz-Beltra et al., 2009] R. Lahoz-Beltra, G. Ochoa, U. Aickelin. 2009. Cheating for Problem Solving: A Genetic Algorithm with Social Interactions. In: Proceedings of Genetic and Evolutionary Computation Conference (GECCO-09), ACM: 811-817.

[Lahoz-Beltra, 2001] R. Lahoz-Beltra. 2001. Evolving hardware as model of enzyme evolution. BioSystems 61: 15-25.

[Lahoz-Beltra, 2004] R. Lahoz-Beltra. 2004. Bioinformatica: Simulacion, Vida Artificial e Inteligencia Artificial (Transl.: Spanish). Ediciones Díaz de Santos, Madrid, Spain.

[Lahoz-Beltra, 2008] R. Lahoz-Beltra. 2008. ¿Juega Darwin a los Dados? (Transl.: Spanish). Editorial NIVOLA, Madrid, Spain.

[Margulis, 1981] L. Margulis. 1981. Symbiosis in Cell Evolution. W.H.Freeman & Company.

[Ogunseitan, 2008] O.A. Ogunseitan. 2008. Genetic transduction in freshwater ecosystems. Freshwater Biology 53: 1228-1239.

[Perales-Gravan and Lahoz-Beltra, 2004] C. Perales-Gravan, R. Lahoz-Beltra. 2004. Evolving morphogenetic fields in the zebra skin pattern based on Turing's morphogen hypothesis. International Journal of Applied Mathematics and Computer Science 14: 351-361.

[Perales-Gravan and Lahoz-Beltra, 2008] C. Perales-Gravan, R. Lahoz-Beltra. 2008. An AM radio receiver designed with a genetic algorithm based on a bacterial conjugation genetic operator. IEEE Transactions on Evolutionary Computation 12(2): 129-142.

[Swiss Federal Institute of Technology Zurich, 2008] Swiss Federal Institute of Technology Zurich. 2008. Systems Optimization. Test Problem Suite. www.tik.ee.ethz.ch/sop/download/supplementary/testProblemSuite/

[Syvanen, 1985] M. Syvanen. 1985. Cross-species transfer. Implications for a new theory of evolution. Journal of Theoretical Biology 112: 333-343.

[von Neumann, 1966] J. von Neumann. 1966. In: The theory of self–reproducing automata (Ed. A. Burks). University of Illinois Press, Urbana, IL.

Authors' Information

Carlos Perales Gravan *– Contact Center Analysis Manager, Motor Area, Línea Directa Aseguradora, Madrid, Spain; e-mail: soyperales+rsch@gmail.com*

Major Fields of Scientific Research: Genetic Algorithms, Time Series, Contact Center Optimization, Processes Automation, Neuronal Networks, Queueing Theory and Android Development.

Javier de Vicente Buendia *– E.U. de Ingenieria Tecnica Aeronautica. Departamento de Matematica Aplicada y Estadistica, Universidad Politecnica de Madrid, Spain.*

Major Fields of Scientific Research: Numerical Methods and Applications to Aerospace Technology

Juan B. Castellanos Peñuela *– Departamento de Inteligencia Artificial, Facultad de Informatica, Universidad Politecnica de Madrid, Spain; e-mail:jcastellanos@fi.upm.es*

Major Fields of Scientific Research: Natural computing, membrane computing, molecular computing and artificial neural networks.

Rafael Lahoz–Beltra *– Department of Applied Mathematics, Faculty of Biological Sciences, Complutense University of Madrid, 28040 Madrid, Spain ; e-mail: lahozraf@bio.ucm.es*

Major Fields of Scientific Research: evolutionary computation, embryo development modeling and the design of bioinspired algorithms